Jostarnameh Comparative Literature
(Scientific Journal Islamic Azad University of Yazd)
Year one / number two / Winter 2017)

Readability of Translation in Scientific Texts with Long Sentences: A Case of a Genetics Textbook

Naeime Naderi[*], Dr. Ali Akbar Moeen[**], Dr. Sima Sayadian[***]

Abstract:

The issue of readability of long sentences has been remarkably noticed in the translation of scientific texts in the world. This subject also demands more attention in translation from English into Persian to solve the difficulty in comprehension of scientific texts in Iran. Accordingly, this quantitative corpus-based study tended to explore readability in scientific texts with long sentences. Hence, the readability of long sentences of a Genetics book in Medicine was compared against its Persian translation. Comparative stylistic model of Vinay and Darbelnet (1958) was used to explore the procedures which were employed in the translation of long sentences and influenced readability of the sentences. Moreover, Lix formula was employed to calculate the readability of long sentences in both source and target texts. The findings revealed that the most important factors in increasing the readability of the translation were as follows; splitting long sentences, changing phrases into clauses and nouns to verbs that decrease the ratio between the number of nouns and the number of verbs, and changing abstract concept words into concrete ones that are modulations. These findings suggest that long and difficult words need to be used less frequently and that passive sentences are to be changed into active ones.

Key words: Readability, Long Sentences, Scientific Texts, Lix Formula.

---

[*] *Department of English, Yazd Branch, Islamic Azad University, Yazd, ma.naderi@iauyazd.ac.ir.*

[**] *Department of English, Yazd Branch, Islamic Azad University, Yazd, Iran, aamoeen@gmail.com.*

[***] *Department of English, Maybod Branch, Islamic Azad University, Yazd, Iran, SimaSayadian@iauyazd.ac.ir.*

1. Introduction

Translation of long sentences in scientific texts requires more attention to the achievement of the readable texts. An extensive part of translation involves scientific texts which mostly comprise of long sentences (Eunson, 2005). Hence, to achieve the targets of the text such as understanding informative texts easily, there is a need to render short and simple sentences rather than long and complex ones.

Readability is that manner of texts in which readers realize texts are easy or difficult to understand (Ateşman, 1997). Readability is not only about the length of sentences and words. Common advice for writing a readable text is to use simple sentences, to use short and familiar words, to use active voice and present tense (Nyqvist, 2012). The longer sentences often need a reader to spend more time to understand. To increase readability, simple sentences are recommended (Rubens, 2002). In English language, there are four sentence types: first, the simple sentence, which contains one independent clause; second, the compound sentence, which contains two or more independent clauses; third, the complex sentence, which contains one independent clause and one or more dependent clauses; and four, the compound/ complex sentence, which contains two or more independent clauses and one or more dependent clauses (Rubens, 2002). In Persian language, sentences are put into two types: first, the simple sentences which contain one verb and second, the compound sentences which comprise more than one verb; these sentence can consist of independent clauses or dependent and independent clauses (Khayampoor, 2014).

Employing long sentences in texts can be indicating of an author's high potentiality in thinking and processing of various paragraphs and dependent clauses of the sentences; these sentences point to the connections among the writer's opinions. Therefore, it avoids monotony in the text and indicates the sympathy of the paragraphs. scholars have considered various numbers of words for long sentences; more than 19 words (Oliveira, Wong, & Hong, 2010), 20 words (Nyqvist, 2012), 25 words (Roh, Hong, Choi, Lee, & Park, 2003), and 50 words (Gerber & Hovy, 1998). Nevertheless, the present research study similar to niqvist, 2012 considered English long sentences as the sentences which had more than 20 words.

Translators of scientific texts need to be aware of that scientific information and also the written principles of both source and target

languages. since the sentence length is regarded to influence the level of difficulty of a text, style manuals and writing guides for informative texts often recommend sentences to be short and simple (Rubens, 2002).

To the researcher's best knowledge, few research studies have to date focused on the translation of long sentences from English into Persian. This is while the prominent differences between Persian and English as source and target languages may affect the translations. Moreover, Scientific texts are written for practical and actual practices and can be applicable to the real life of people. These issues require more attention because the low quality of translation of scientific texts causes the readers to lose interest in reading and studying such texts. Furthermore, Persian readers usually prefer to study the scientific texts with active tenses, small words and comprehensible sentences without rereading to comprehend the original meaning. Therefore, providing a good translation in the scientific texts is very vital.

The results of this research could help translators, translation students, and teachers in translating. Translators of scientific texts are usually dealing with the problems of complicated structures of English long sentences. Hence, employing the results of this research can be beneficial to them.

The research questions of the present study are as follows:

1) What is the frequency of long sentences and long words in the English Genetics text?

2) What is the frequency of long sentences and long words of the translated text into Persian?

3) Is there any significant difference between the number of words of long sentences in English and Persian?

4) What differences exist between English and Persian sentences based on Vinay and Darbelnet's comparative stylistic model?

5) How much readability do both the original text and the translated long sentences from English to Persian have based on Lix formula?

2. Review of the Related Literature

2-1. Readability

Readability is the quality of the text when readers can understand that easily and enjoyable. As (Anderson, 1981) indicated, Lix is a readability formula from Sweden which assessed the difficulty of texts in different languages like English. The Sweden Carl-Hugo Bjornson (1968) developed

the Lix formula and called it 'Lasbarhets index' and Lix is the shortened form of its name. To calculate Lix formula, the average number of words in a long sentence is added to the percent of words of more than six letters (Anderson, 1981).

As Anderson (1981) indicated, for the calculation of readability, a calculated value less than 30 corresponds to children's books that are very easy to read. If the value is between 30 and 40, the text is regarded as easy to read, for instance, fiction or popular magazines. If it is between 40 and 50, the text is of average difficulty, corresponding to normal newspaper text. If it is between 50 and 60, the text is regarded difficult and it is a typical value of official texts. If the value exceeds 60, it is regarded as very difficult.

Anderson (1981) argued about Lix as an easy formula to compute (assuring interceder reliability) and useful for examining texts at various levels, whose material was allocated to people at different ages, and also across English and non-English languages. Moreover, he in his seminal paper (1983) introduced Lix and Rix, a quicker index of lix, as variations on a little-known readability index and the easy scores that teachers and librarian used to determine the reading difficulty of English and foreign language materials. He in his both papers, tried to introduce and emphasize Lix as a useful index to examine the readability of different levels of English and non-English texts, although he also in the second research introduced another index of Lix, Rix, as a readability index.

Kondru (2007) proposed a new readability formula and employed a part-of-speech structure, to compare this formula with other formulas. Their experimental results, in contrast to previous formulas, showed that the new formula led to more noticeable results than other readability formulas. In another research, Izgi and Seker (2012) compared various readability formulas to examine the readability value of some science textbooks of two grades. Exploring this readability value showed that the outputs were different from one another. In another paper, Chiang, Englebrecht, Phillips Jr, and Wang (2008) investigated whether there were readability differences among leading principles of accounting textbooks. Seven financial accounting principles textbooks were compared using four readability evaluation methods. Evidence suggested that readability among textbooks varied, but readability within textbooks was generally consistent.

Although all three research studies compared readability of each text by using different readability formulae, Kondru's (2007) research presented a

new formula with better grade of readability index in contrast to the other researchers. In the other words, Izgi and Seker (2012) and Chiang, et al. (2008) compared various readability formulas to examine the text readability value, but they did not present a new readability formula.

Mavasoglu and Dincer (2014) analyzed French language teaching websites and textbooks in terms of their lexical density and clause length which were common criteria for readability. Results showed that the websites, with their high number of words and clauses, are more likely to present a difficult readability index than the textbooks. Weir and Anagnostou (2008) were the researchers who rather than employing word and clause or sentence length, used collocation frequency as a readability factor since the readability of any text would be affected adversely by the presence of collocations. So this study was different from the two previous ones.

Remus (2011) used 6 different readability formulas to improve subjective classification in the level of the sentences. He demonstrated that employing readability formulas and their compound as traits could significantly improve the subjective quality of classifying sentences. But Dell'Orletta, Wieling, Venturi, Cimino, and Montemagni (2014) investigated the problem of sentence readability assessment and addressed two common problems in relation to text simplification, that was the use of corpus in order to train, and the recognition of the worthwhile traits to define readability of sentences. They concluded that determination of the sentence readability was not simple and that it needed many traits, primarily syntaxes. Although both researchers, Remus (2011) and Dell'Orletta et al. (2014) investigated the problem of sentence length, their purposes and procedures were totally different.

Alotaibi, Alyahya, Al-Khalifa, Alageel, and Abanmy (2016), Chi, Jabbour, and Aaronson (2017), Lynch et al. (2017) and Tran, Singh, Singhal, Rudd, and Lee (2017) measured the readability of medical information on leaflets, websites and online resources. The researchers showed that the above materials were too sophisticated for the average adults whose literary skills were well documented.

2-2. Theoretical Framework

The classical *Comparative Stylistic* model of Vinay and Darbelnet (1958) was employed to compare the differences between the English and

the Persian Genetics textbooks. The theory involves two common strategies of translation; oblique and direct translation.

Direct translation involves three procedures – borrowing, directly transference of SL word to TL; calque, transference of SL expression or structure in the word for word translation; and literal or word-for-word translation (Vinay & Darbelnet, 1995). Oblique translation comprises of four procedures: transposition, modulation, equivalence, and adaptation. 1) Transposition is a common structural change and is probably most frequently used by translators. Transposition implies a shift from a part of speech to another, but the sense does not change. e.g. a verb is changed into a noun, and an adverb is changed into a verb (Fawcett, 2014). 2) Modulation shifts the meaning and point of view of the source text (ST). This shift could be reasonable if, either a word-for-word or even oblique, translation causes in a structure with correct grammar, that is inappropriate, difficult or unskilled in the target text (TT). In this procedure, "an abstract word is changed into a concrete one, active voice is changed into passive, and a cause is changed into an effect, whole to part, part to another part, active to passive, and also a reversal of terms, a negation of opposite, rethinking of intervals and limits in space and time and change of symbol" (Munday, 2016, p. 88).

According to Munday (2016), several supplementary translation procedures were defined by Vinay and Darbelnet (1958) including amplification and economy. In amplification, the TL uses more words, often because of syntactic expansion, but the opposite of amplification is an economy.

3. Material and Method

This corpus-based study adopted a quantitative approach entailing a descriptive and evaluative analysis. In the evaluative section of the study, the researcher applied a formula, Lix, to compare readability in source and target text to come up with representative results.

The comparable corpora which was employed in this study was seventh edition of a scientific book, by Nussbaum, McInnes, and Willard (2007) which was composed of 20 chapters and about 600 pages. Moreover, its translation by Mohammad Ghofrani, Majid Zaki and Lale Habibi (1388) consisted of 20 chapters and about 800 pages. The main reason for selecting

this book as the corpus of the study was possessing a large number of long sentences in a scientific text.

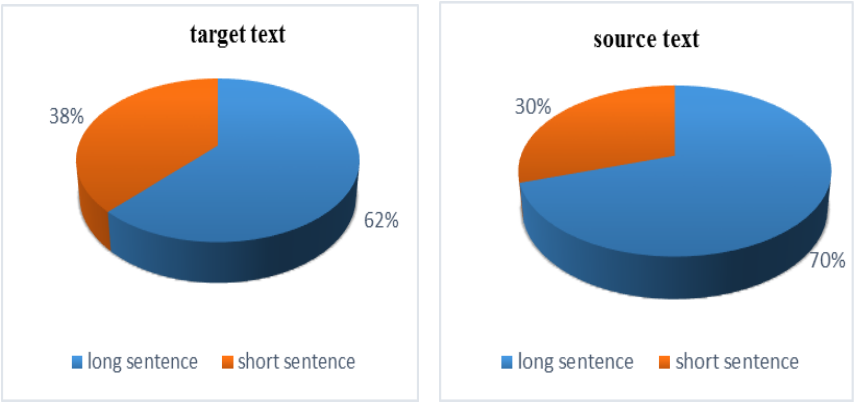
The researcher identified long sentences (more than 20 words) in both source text (Genetics in Medicine, 2007) and target text (its translation in Persian). In the second step, the number of long words (more than 6 letters) of each long sentence was counted. Th en, readability of the original text using the number of words and long words of a long sentence and the readability Lix index was calculated (average sentence length added to the percentage of long words in the text). Then, using Vinay and Darbelnet's (1958) 'comparative stylistic model', the differences between two texts from two languages were considered and different procedures were identified.

4. Analysis

The source text was a scientific book, i.e., Genetics in Medicine. The number of 140 long sentences from the first three pages of four chapters, of the source and target texts - chapters 1, 2, 10 and 20- were identified and analyzed.

4-1. Results Related to Research Questions One, Two and Three

The following figures 1 and 2 are provided to compare long and short sentences in ST and TT.



Figures 1& 2. Percentage of long and short sentences in ST and TT

As shown in figure 1, in ST 62% of sentences were long and 28% were short, but the ratio in target text changed and as seen in figure 2, 70% of

sentences of TT were long and 30% were short. It means that some long sentences were translated into short sentences.

Afterwards, long words (more than 6 letters) were analyzed. The results of collected data related to long words, the percentage of long words in ST and TT and the averages of word length in ST and TT is shown in figure 3.

In ST, among 4820 words of the long sentences, as shown in Figure , 1624 words were long. The TT sentences consisted of 4708 words in which 1428 words were long. In other words, 33 percent of the words in ST and 30 percent of the words in TT were long. In average, every sentence of ST included about 11.6 long words, but in the TT the contribution of long words of each sentence was 9.5. It leads to this result that the number of long words of TT in average has rather decreased.

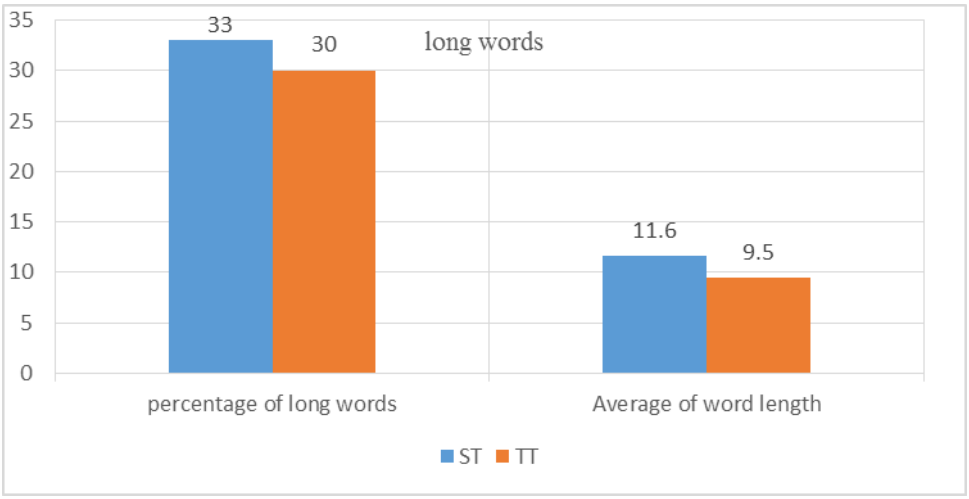


Figure 3. Percentage of long words in the ST and TT

4-2. Results Related to Research Questions 4 and 5

Vinay and Darbernet's (1958) comparative stylistic framework was applied to compare the long sentences of both source and target texts. The comparison of sentences based on the framework showed the percentage of applied procedure in the translation of long sentences.

As shown in figure 4, the comparison of English text and the Persian translation led to this result that some procedures had a prominent role in the translation of this scientific text: economy (29%), modulation (26%), amplification (17%) and borrowing (13%). Some other procedures were less vigorously effective, such as transposition (7.8%), literal translation (4.7%)

and calque (1.4%). But a less obvious role was seen in the other procedures related to this framework.

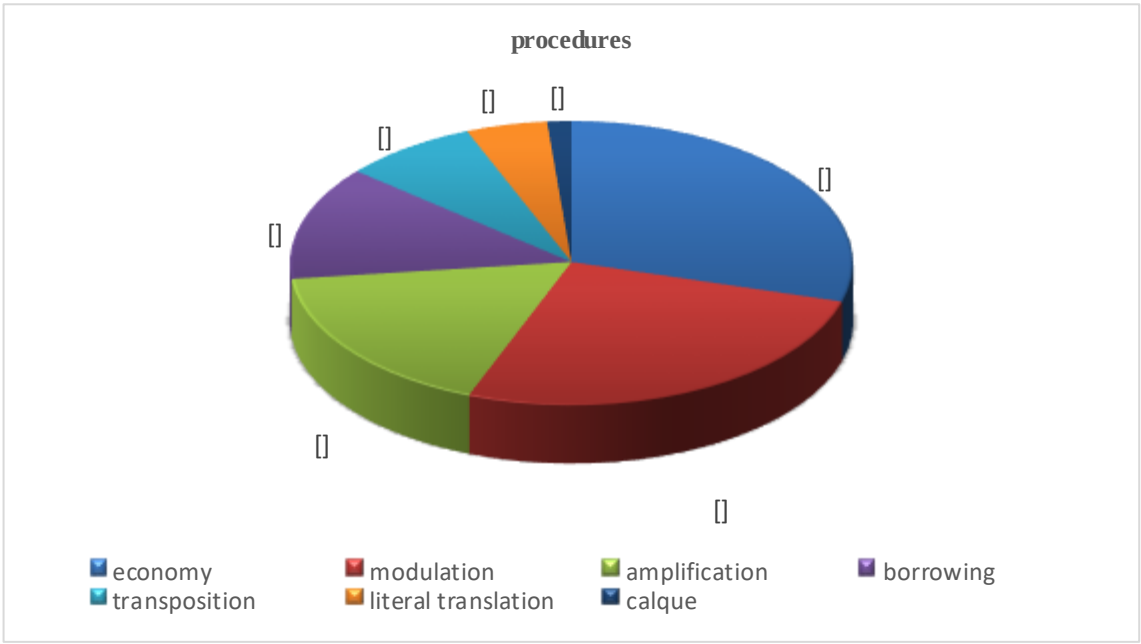


Figure 4. Percentage of applied procedures of comparative stylistic model in translated text

To measure the readability of long sentences in source and target texts, Lix index was used. To calculate the Lix value, the number of words and long words of long sentences were necessary (table 4.1). As we know, in writing Persian words, short vowels are usually not written, while in English texts almost all vowels are written. Therefore, in order to provide equal circumstances for counting the letters of Persian long words, vowels were considered. For example, the equivalence of medicine in Persian is /*peze∕ki*/ which was considered to consist of 7 letters.

The average Lix scores of both the original and translation texts are shown in figure 5, in below.

As shown in figure 5, the average Lix score of both source and target text is high and according to Lix tools for calculating readability if the calculated value of the text exceeds 60, it is regarded as very difficult (Anderson, 1981). Yet, Lix scores of the source text were somewhat higher than the target text. Therefore, the original text was more difficult to read than the translated text.

Figure 5. The average Lix scores in the ST and TT

4-3. Samples

The readability Lix index of the Persian translation of some long sentences was almost similar to their English sentences. While some other Persian sentences earned higher readability Lix values or lower values rather than their original sentences in the source text.

Sample One

Sentence one has gained almost the similar readability values in the source and target texts.

Genetics in medicine had its start at the beginning of the 20th century, with the recognition by Garrod and others that Mendel's laws of inheritance could explain the recurrence of certain disorders in families.

p.1, in ST

نقش ژنتیک درپزشکی ازاوایل قرن بیستم وبا کشف این واقعیت توسط گارود و سایرین که قوانین مندل برای توارث می‌توانند علت بروز مجدد برخی بیماری‌ها را در خانواده‌ها توضیح دهند، شناخته شد.

p.23, in TT

Sentence 1, in the ST, was a complex sentence which included 34 letters and 12 long words. The TT also earned almost similar characteristics; one sentence with 35 letters and 12 long words. The long sentence comprised of one dependent and one independent clause. The subject of the long sentence

in ST is *genetics in medicine* and in TT is نقـش ژنتیـک درپزشـکی. The comparison between two sentences showed that three noun phrases were added to the Persian translation sentence: نقـش, علـت and ایـن واقعیـت. These added phrases were regarded as an amplification procedure according to the comparative stylistic model of Vinay and Darbelnet (1958). The verb phrase *had* was translated into شـناخته شـد. The change concerned a modulation procedure; because the point of view of the ST has changed. Another point related to the translation of this sentence was about borrowing the noun *genetics* from English to Persian. Moreover, the ratio between the number of nouns and the number of verbs decreased from 7 to 6.3.

To sum up, the English sentence 1 and its equivalence in Persian were almost similar in the number of clauses, verbs, nouns. Furthermore, the number of words and long words were almost the same in ST and TT. That was why the two long sentences earned almost the similar Lix values (the English sentence, 69.3 and the Persian sentence, 69.2). Since the Lix value >60 shows the text is very difficult (Anderson, 1981), these sentences in ST and TT were regarded as very difficult.

Sample Two

In the following complex sentence, the readability Lix value of sentences in the ST and TT received many different levels and the readability of the Persian sentence increased noticeably.

Medical genetics has become part of the broader field of genomic medicine, which seeks to apply a large-scale analysis of the human genome, including the control of gene expression, human gene variation, and interactions between genes and the environment, to improve medical care.

p.1, in ST

ژنتیک پزشکی به جزیـی از یـک محـدودهٔ بزرگ‌تر به نام پزشکی ژنومی تبدیل شده اسـت. هدف پزشکی ژنومی کـاربردی کـردن تجزیـه و تحلیل ژنوم انسان در مقیاس بالا، مانند کنتـرل بیـان ژن؛ تنوعات ژن‌های انسان؛ میانکنش بین ژن‌ها و محیط، با هدف بهبود مراقبت‌های پزشکی است.

p.23, in TT

In ST, the above sentence was complex and consisted of 38 words and 16 long words. In TT, this sentence split into sentences with 15 and 31

words. According to the definition of long sentences, the first sentence was short and the second was long (more than 20 words).

The complex sentence in ST was comprised of one main clause and two dependent clauses. The subject of the main clause was the noun phrase: *medical genetics* and the subject of the independent clause was a conjunction: *which* that referred to 'medical genetics'. But, the Persian text was comprised of one short and one long sentence. The subject of the first sentence of the TT was ژنتیک پزشکی and the second sentence was هدف پزشکی ژنومی .

There were some changes based on the comparative stylistic model of Vinay and Darbelnet (1958); the verb *seek* was translated into the noun هدف and the verb *improve* was rendered to the noun بهبـود . The changes were transposition or changes from one part of speech to another.

Since one English sentence was translated into two Persian sentences, the conjunction *which* was deleted and the noun هـدف was added to the start of the second sentence. According to the comparative model of Vinay and Darblenet (1958), the change concerning deletion was an economy and the second shift was amplification. The other procedures applied in translated texts were related to borrowing the noun *genetics* and literal translation of part of the sentence. The ratio between the number of nouns and the number of verbs increased from 4.5 in ST to 12 in TT.

To sum up, splitting the complex sentence into two sentences one of which was short was the important reason to decrease the readability index of the translated sentence from very difficult to medium (from 80.1 to 49). Therefore, the readability of the translated sentence increased.

Sample Three

In the following example, the readability of Persian sentence decreased:

A hematologist combines family and medical history with gene testing of a young adult with deep venous thrombosis to assess the benefits and risks of initiating and maintaining anticoagulant therapy.

p.2, in the ST

یک متخصص خون سابقهٔ خانوادگی، تاریخچـهٔ پزشکی و نتایج آزمایش ژنتیکی یک جوان مبـتلا به خون‌لختگی (ترومبوز) وریـد عمقـی را مـورد مطالعه قرار می‌دهد تا منـافع و مضـرات آغـاز و ادامهٔ درمان ضدانعقادی را برای او ارزیابی کند.

p.24, in the TT

This English sentence was complex and consisted of 30 words and 11 long words; it was comprised of one independent clause whose subject was the noun phrase: *A hematologist* and one dependent clause whose subject referred to the previous clause. But the Persian sentence was a complex sentence with 38 words and 16 long words. It was comprised of the main clause whose subject was the noun phrase: یـک متخصـص خـون and a dependent clause whose subject was the subject of the previous clause.

The employed procedures in the Persian sentence included borrowing *thrombosis* which was translated ترومبـوز, calque or literal translation of expression *anticoagulant* into ضد انعقـادی, amplification or using more words like تاریخچـه and نتـایج in the independent clause and modulation or changing the meaning of the verb co*mbines* to مورد مطالعه قرار می‌دهد.

The ratio between the number of nouns and the number of verbs increased from 7 to 10.5. The readability Lix value of this sentence in the ST was 66.6 and in the TT was 80.1.

In sum, although the readability of both English and Persian sentences was regarded as very difficult (>60), the readability of the Persian sentence considerably increased. In relation to readability, some procedures including borrowing and calque had an important role because of transferring difficult and long words in the target text, as well as amplification and modulation. In other words, one word was translated into several words. All these mentioned features led to a Persian sentence with more words and longer words. Therefore, the readability of the sentence decreased.

Sample Four

In the following sentence, the readability of translated sentence increased considerably:

The paternal and maternal homologs exchange homologous segments by crossing over and creating new chromosomes that are a "patchwork" consisting of alternating portions of the grandmother's chromosomes and the grandfather's chromosomes.

p.208, in the ST

همولوگ‌هـای پـدری و مـادری، قطعـات همولـوگ خودشـان را توسـط تبـادل متقـاطع تعویض می‌کنند و کروموزوم های جدیدی ایجاد می‌کنند که به‌صورت "وصله وصله" می‌باشـند و از قطعات متناوب کروموزوم هـای پـدربزرگ و مادربزرگ تشکیل شده‌اند.

p.304, in the TT

The English sentence was complex and consisted of 31 words and 17 long words. But, the Persian sentence included 33 words and 6 long words. The sentence in ST was comprised of two clauses, one independent and two dependent clauses, the subject of the sentence was *The paternal and maternal homologs*. While the sentence in the TT was compound- complex and it was comprised of three main clauses and one dependent clause: کـه بـه صورت "وصله وصله می باشند.

The noun phrase *creating new chromosomes* was translated into the independent clause کروموزوم های جدیدی ایجاد می کننـد. The dependent clause *consisting of alternating portions of the grandmother's chromosomes and the grandfather's chromosomes* was translated into the independent clause از قطعات متناوب کروموزوم های پدربزرگ ومادربزرگ تشکیل شده انـد. Both changes showed transposition or changing part of speech, based on the comparative stylistic framework of Vinay and Darbelnet (1958).

In the ST, the number of nouns was 13 and the number of verbs was 2, so the ratio between nouns and verbs was 6.5. In the TT, the number of nouns was 11 and the number of verbs was 4, so the ratio between nouns and verbs was 2.7. The calculated Lix value in the ST was 85.8 and in the TT was 51.1

To sum up, changing a phrase and a clause to independent clauses led to the change of some nouns to verbs. Then the ratio between nouns and verbs decreased (from 6.5 in the ST to 2.7 in the TT). Moreover, the number of long words decreased from 17 to 6. All these mentioned features noticeably increased the readability of the translated long sentence. In other words, the readability of Persian sentence changed from very difficult in the ST into difficult in the TT.

Sample Five

In the following sentence, the readability of long sentence of the target text was reduced:

One, two, or more recombinations occurring between two loci at the

یک، دو یا تعداد بیشتری از تبـادل متقـاطع بین دو جایگاه در مرحله چهار کروماتیدی منجر

four-chromatid stage result in gametes that are 50% nonrecombinant (parental) and 50% recombinant (nonparental), which is precisely the same proportions one sees if the loci were on different chromosomes.

p.208 in the ST

به تولید گامت‌هایی می‌گردد که ۵۰ درصد غیرنوترکیب (والدینی) و ۵۰ درصد نوترکیب (غیروالدینی) هستند؛ این نسبت دقیقاً مشابه نسبت‌های حالتی است که دو جایگاه در روی دو کروموزوم مجزا واقع هستند.

p.304 in the TT

In the ST, the sentence was complex and was consisted of 40 words and 13 long words. The Persian sentence was also complex and included 49 words and 16 long words. Both the English and Persian sentences were comprised of one main clause and 5 dependent clauses.

The phrase *occurring* and the clause *one sees* were limited, it is an economy, according to Vinay and Darblenet (1958). The noun phrase این نسبت was added to the Persian sentence, i.e. amplification and the noun phrase *recombinations* was translated into the different meaning, i.e. modulation: تبادل متقاطع . Changing the clause *one, two, or more recombinations occurring* into the phrase یک، دو یا تعداد بیشتری از تبادل متقاطع is a transposition.

In ST, the number of nouns was 9 and the number of verbs was 5. In the TT, the number of nouns was 14 and the number of verbs was 4. The ratio between the number of nouns and the number of verbs increased from 1.8 to 3.5. The readability Lix value of the English sentence was72.5 and the Persian sentence was 81.5. Although, the readability of both sentences was regarded as very difficult, the readability of the Persian sentence reduced.

To sum up, changing a clause to a phrase, deletion of a clause and amplification of phrases in the Persian sentence increased the number of words and long words and decreased the readability of Persian complex sentence.

5. Discussion

The present study followed Niqvist's (2012) research study. The strategies which she used to make the text more readable were as follows: Transpositions, for example, nouns were changed into verbs, and in this way long noun phrases were changed into clauses with active verbs; Modulations,

especially abstract nouns were changed into more concrete ones and, sometimes, the perspective changed when finding new subjects to the divided sentences. Hence, the results were parallel; splitting long sentences into two or more shorter ones, applying short and familiar words, and simple sentences, changing nouns to verbs and noun phrases to clauses, using modulations –especially explicative, abstract nouns and active tenses – were the results which were attained in both studies. However, the target text related to this study was a Persian scientific textbook, which differed from the Swedish lecture that Niqvist (2012) analyzed. Moreover, this study considered not only seven procedures of the comparative stylistic model as Niqvist (2012) applied, but also supplementary translation procedures of Vinay and Darbelnet (1958).

The results of this study also supported some results of the study of Lei and Zhang (2010) in achieving the information equivalence in EST translation of long sentences such as splitting, embedding and synthetic approach of translation. But there was no idea about the other results of EST translation like interlinear translation, inverse translation. The present study confirmed Brown, Della Pietra, Della Pietra, Mercer, and Mohanty's (1992) research finding that breaking down long words into pieces increased readability, but less than 20 words not 10 words.

This study supported Oliveira et al. (2010), Kim, Zhang, and Kim (2001) and Gerber and Hovy, (1998) in breaking down the length of the sentences to make the lower sentences which were easy to understand.

The findings of the present research confirmed Anderson's (1981) argument about Lix. He argued that Lix formula could examine texts across English and non-English languages. So, this study, as it explored the Persian text, confirmed his claim.

Smith and Jönsson (2011) used various measures including Lix, NR and (OVIX) to measure the amount of vocabulary, opinions, curiosity, and structure of sentences in the text. Their results showed that the summarized texts were more readable. The present study differed in many aspects; 1) the original text was compared with the translated text rather than the summary of the text, 2) the study used only Lix as the measurement tool and 3) the subjects of measure differed and they were only in sentence structure. However, the results were similar in increasing readability in the second texts.

The procedures and the framework, Vinay and Darbelnet's comparative stylistic model, of this study were in contrast with three following research studies. Kondru (2007) proposed a new readability formula and worked on the part of speech of the sentence structures in the texts, compared it with other formulas. Izgi and Seker (2012) compared various readability formulas to determine the readability levels of textbooks. Chiang et al. (2008) investigated whether there were readability differences among leading principles of accounting textbooks. Seven financial accounting principles were compared using four readability evaluation methods.

The objectives of this study were not parallel with the objective, procedures, and conclusion of Leonga, Ewingb, and Pittb (2002) who examined readability of the text on business-to-business (B2B) websites and demonstrated how the relative readability of a company's web pages could be evaluated against those of its competitors. They explored how readability formulae could provide quick feedback to web marketers based on key indicators such as sentence length, proportion of difficult words and frequency of polysyllabic words.

Finally, the objective and conclusion of the present study were relatively similar to the research study of Remus (2011) who assessed the readability of sentence length; however, Remus (2011) used 6 different readability formulas.

6. Conclusion

The purpose of this research was the study of readability in scientific texts with long sentences. In doing so, the comparative stylistic model of Vinay and Darbelnet (1958) and Lix formula were employed. Due to the existence of a large number of long sentences in scientific texts, research on the readability of long sentences seemed necessary.

The results in response to research questions one, two and three showed that the number of long sentences reduced from 70% in ST to 62% in TT (figure 1 & 2). It means that some long sentences were translated into short ones. Moreover, the number of long words of TT in average had rather decreased, from 33% in ST to 30% in TT (figure 3).

Calculating readability Lix formula for each sentence showed that the average Lix scores of long sentences reduced from 67.4 in ST to 62.8 in TT. As sentences with readability Lix score >60 are regarded as very difficult

(Anderson, 1981), the English and Persian long sentences with respectively Lix scores between 67.4 and 62.8 were regarded as very difficult. The gathered data, in response to research question 5, revealed that in the Persian text, like the English text, most of the sentences were regarded as difficult (50- 60) and very difficult to read (> 60).

Analyzing sentences based on the comparative stylistic model of Vinay and Darbelnet (1958), in response to research question 4 showed that some procedures and factors had an important role in dividing readability of long sentences into three categories. In the first category, some factors increased the readability of long sentences; those sentences which split into 2 or more short sentences were more readable than the original sentences. Moreover, some changes in semantics of the ST including changing abstract concept words to concrete concept ones, changing passive voice to active voice and using explicative phrases which were easy to understand, i.e. modulation resulted in understandable translated sentences compared to their English counterparts. Moreover, using more words to make the sentence more understandable, that is amplification, presenting the main brief information, because the large number of words were deleted from the target sentence i.e. economy and decreasing the number of long words increased the readability of translated sentences. Furthermore, changing phrases into clauses, i.e. transposition which led to changing some nouns to verbs and decreasing the ratio between the number of nouns and verbs made the sentences easy to understand. All these mentioned factors increased the readability of long sentences.

The second category included the original and translated sentences with relatively similar readability, because of some procedures like literal translation, borrowing phrases from the English text, transferring some expressions in the literal translation, i.e. calque, and exiting relatively similar number of words and long words in ST and TT which resulted in approximately similar readability values in English and Persian sentences.

In the third category, the readability of long sentences decreased because of the explications of phrases, using more words and long words as well as increasing the ratio between the number of nouns and verbs in the Persian sentence which caused a decrease in the readability of translated sentences. Moreover, the changes of clauses to phrases, deletion of clauses and amplification of phrases in the Persian sentence which caused an

increase in the number of words and long words, decreased the readability of Persian long sentence.

At last, amplification which included the use of more words to elucidate the Persian sentence, increased the readability of target sentences which were split into two or more sentences. However, since amplification increases the number of words in a long sentence, the readability of a long sentence was reduced.

The theoretical implications of the study indicated that since few research studies have concentrated on the issue of readability of long sentences in scientific texts, more studies of this kind seem necessary. Moreover, many students and teachers deal with the problem of reading translated textbooks with long sentences which were difficult or very difficult to understand. These unreadable long sentences may transfer to textbooks and scientific research studies in which these texts are used as references and provide gaps in the knowledge of their readers. Hence, it is essential to study how long sentences can be translated into more readable sentences.

The most fruitful area of further research would be to investigate readability of two or more translations of a scientific text. It would be extremely interesting to study the readability of long sentences through readers' opinions. Readers can include translation students, professional translators or readers of the related scientific text. Also, more investigation can be conducted on comparing the calculated results of other readability formulae.

7. References

1. Alotaibi, S., Alyahya, M., Al-Khalifa, H., Alageel, S., & Abanmy, N. (2016). Readability of arabic medicine information leaflets: A machine learning approach. Procedia Computer Science, 82, 122-126.
2. Anderson, J. (1981). Analysing the readability of English and non-English texts in the classroom with Lix. Paper presented at the Annual Meeting of the Australian Reading-Association, Darwin, Australia.
3. Ateşman, E. (1997). Measuring readability in Turkish. AU Tömer Language Journal, 58, 171-174.
4. Brown, P. F., Della Pietra, S. A., Della Pietra, V. J., Mercer, R. L., & Mohanty, S. (1992). Dividing and conquering long sentences in a translation system. Paper presented at the Proceedings of the workshop on Speech and Natural Language, Yorktown Heights.

5. Chi, E., Jabbour, N., & Aaronson, N. L. (2017). Quality and readability of websites for patient information on tonsillectomy and sleep apnea. International Journal of Pediatric Otorhinolaryngology, 98, 1-3.
6. Chiang, W.-C., Englebrecht, T. D., Phillips Jr, T. J., & Wang, Y. (2008). Readability of financial accounting principles textbooks. The Accounting Educators' Journal, 18, 47-80.
7. Dell'Orletta, F., Wieling, M., Venturi, G., Cimino, A., & Montemagni, S. (2014). Assessing the readability of sentences: Which corpora and features? Paper presented at the Proceedings of the Ninth Workshop on Innovative Use of NLP for Building Educational Applications, Baltimore, Maryland USA.
8. Eunson, B. (2005). Communicating in the 21st century. Australia: : John Wiley & Sons.
9. Fawcett, P. (2014). Translation and language. London: Routledge.
10. Gerber, L., & Hovy, E. (1998). Improving translation quality by manipulating sentence length. Paper presented at the Conference of the Association for Machine Translation, , The Americas.
11. Izgi, U., & Seker, B. S. (2012). Comparing different readability formulas on the examples of science-technology and social science textbooks. Procedia-Social and Behavioral Sciences, 46, 178-182.
12. Khayampoor, A. (2014). Persian grammar (15th ed.). Tabriz: Sotoode.
13. Kim, S.-D., Zhang, B.-T., & Kim, Y. T. (2001). Learning-based intrasentence segmentation for efficient translation of long sentences. Machine Translation, 16(3), 151-174.
14. Kondru, J. (2007). Using part of speech structure of text in the prediction of its readability (master thesis), University of Texas, Arlington, University of Texas, Arlington.
15. Lei, X.-f., & Zhang, W.-t. (2010). Exploring the achievement of information equivalence in EST translation of long sentences Paper presented at the National Teaching Seminar on Cryptography and Information Security, Kunming, China.
16. Leonga, E. K. F., Ewingb, M. T., & Pittb, L. F. (2002). E-comprehension: Evaluating B2B websites using readability formulae. Industrial Marketing Management, 31(2), 125-131.
17. Lynch, N. P., Lang, B., Angelov, S., McGarrigle, S. A., Boyle, T. J., Al-Azawi, D., & Connolly, E. M. (2017). Breast reconstruction post mastectomy-Let's Google it. Accessibility, readability and quality of online information. The Breast, 32, 126-129.
18. Mavasoglu, M., & Dincer, S. (2014). Readability and french language teaching texts: an analysis of french language teaching websites and textbooks. Procedia-Social and Behavioral Sciences, 116, 256-259.
19. Munday, J. (2016). Introducing translation studies: Theories and

applications. London: Routledge.
20. Nussbaum, R. L., McInnes, R. R., & Willard, H. F. (2007). Thompson & Thompson genetics in medicine E-Book: Elsevier Health Sciences.
21. Nyqvist, A. (2012). To translate and adapt a text with long sentences: With focus on readability. master thesis. Linnaeus University, Sweden.
22. Oliveira, F., Wong, F., & Hong, I.-S. (2010). Systematic processing of long sentences in rule based Portuguese-Chinese machine translation. Paper presented at the International Conference on Intelligent Text Processing and Computational Linguistics, Iaşi, Romania.
23. Remus, R. (2011). Improving sentence-level subjectivity classification through readability measurement. Paper presented at the 18th Nordic Conference of Computational Linguistics, Riga, Latvia.
24. Roh, Y.-H., Hong, M., Choi, S.-K., Lee, K.-Y., & Park, S.-K. (2003). For the proper treatment of long sentences in a sentence pattern-based English-Korean MT system. Paper presented at the MT Summit IX, New Orleans , USA.
25. Rottensteiner, S. (2010). Structure, function and readability of new textbooks in relation to comprehension. Procedia-Social and Behavioral Sciences, 2(2), 3892-3898.
26. Rubens, P. (2002). Science and technical writing: A manual of style. London: Routledge.
27. Smith, C., & Jönsson, A. (2011). Automatic summarization as means of simplifying texts, an evaluation for swedish. Paper presented at the 18th Nordic Conference of Computational Linguistics, Riga, Latvia
28. Tran, B. N. N., Singh, M., Singhal, D., Rudd, R., & Lee, B. T. (2017). Readability, complexity, and suitability of online resources for mastectomy and lumpectomy. Journal of Surgical Research, 212, 214-221.
29. Vinay, J.-P., & Darbelnet, J. (1995). Comparative stylistics of French and English: a methodology for translation (Vol. 11). Amsterdam, The Netherlands: John Benjamins Publishing.
30. Weir, G., & Anagnostou, N. (2008). Collocation frequency as a readability factor. Paper presented at the Proceedings of the 13th Conference of the Pan Pacific Association of Applied Linguistics, Pattaya, Thailand.